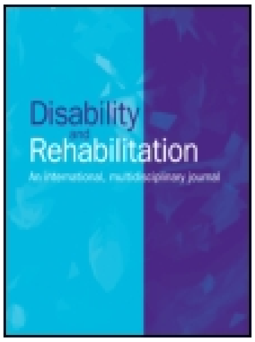

Taylor & Francis
Taylor & Francis Group

Disability and Rehabilitation

ISSN: 0963-8288 (Print) 1464-5165 (Online) Journal homepage: http://www.tandfonline.com/loi/idre20

Reliability and validity of the Falls Efficacy Scale-International after hip fracture in patients aged ≥65 years

Jan H. M. Visschedijk, Caroline B. Terwee, Monique A. A. Caljouw, Monica Spruit-van Eijk, Romke van Balen & Wilco P. Achterberg

To cite this article: Jan H. M. Visschedijk, Caroline B. Terwee, Monique A. A. Caljouw, Monica Spruit-van Eijk, Romke van Balen & Wilco P. Achterberg (2015) Reliability and validity of the Falls Efficacy Scale-International after hip fracture in patients aged ≥65 years, Disability and Rehabilitation, 37:23, 2225-2232

To link to this article: http://dx.doi.org/10.3109/09638288.2014.1002573

Published online: 14 Jan 2015.

Disability and Rehabilitation
An international, multidisciplinary journal

http://informahealthcare.com/dre
ISSN 0963-8288 print/ISSN 1464-5165 online

Disabil Rehabil, 2015; 37(23): 2225–2232
 DOI: 10.3109/09638288.2014.1002573

informa healthcare

ASSESSMENT PROCEDURES

Reliability and validity of the Falls Efficacy Scale-International after hip fracture in patients aged ≥65 years

Jan H. M. Visschedijk[1,2], Caroline B. Terwee[3], Monique A. A. Caljouw[1], Monica Spruit-van Eijk[1], Romke van Balen[1], and Wilco P. Achterberg[1]

[1]*Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, The Netherlands,* [2]*Zorggroep Solis, Deventer, The Netherlands, and* [3]*Department of Epidemiology and Biostatistics and the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands*

Abstract

Purpose: To assess the measurement properties of the Falls Efficacy Scale-International (FES-I) in patients after a hip fracture aged ≥ 65 years. *Methods:* In a sample of 100 patients, we examined the structural validity, internal consistency and construct validity. For the structural validity a confirmatory factor analysis was carried out. For construct validity predetermined hypotheses were tested. In a second sample of 21 older patients the inter-rater reliability was evaluated. *Results:* The factor analysis yielded strong evidence that the FES-I is uni-dimensional in patients with a hip fracture; the Cronbach's alpha was 0.94. When testing the reliability, the intra-class correlation coefficient was 0.72, while the Standard Error of Measurement was 6.4 and the Smallest Detectable Change was 17.7 (on a scale from 16 to 64). The Spearman correlation of the FES-I with the one-item fear of falling instrument was high ($r = 0.68$). The correlation was moderate with instruments measuring functional performance constructs and low with instruments measuring psychological constructs. *Conclusions:* Reliability and structural validity of the FES-I in patients after a hip fracture are good. The construct validity appears more closely related to functional performance constructs than to psychological constructs, suggesting that the concept measured by the FES-I may not capture all aspects of fear of falling.

Keywords

Falls Efficacy Scale-International, fear of falling, hip fractures, measurement properties

History

Received 25 January 2014
Revised 12 December 2014
Accepted 22 December 2014
Published online 14 January 2015

➤ Implications for Rehabilitation

- The Falls Efficacy Scale-International (FES-I), which is commonly used to measure fear of falling in community-dwelling older persons, can also be used to assess fear of falling in patients after a hip fracture.
- The reliability and the structural validity of the FES-I for these hip patients are good, whereas the construct validity of the FES-I is not optimal.
- The FES-I may not capture all aspects of fear of falling and may be more closely related to functional performance than to psychological concepts such as anxiety.

Introduction

The annual incidence of patients with hip fractures is expected to grow substantially in the coming decades, i.e. from 1.3 million in 1990 to about 4.5 million in 2050 [1]. Overall mortality is reported to be 20–33% and only a minority of patients recover completely [2,3]. Psychological factors, such as fear of falling, are associated with these unwanted outcomes [4,5]. Fear of falling may even have more impact on functional recovery than pain or depression [4]. Fear of falling results in avoidance of activities and reduces mobility after a hip fracture [6]. Fear of falling is common among older persons (21–85%) [7]; moreover, in older patients after a hip fracture figures as high as 50–65% are reported [8–10].

Fear of falling has initially been regarded as the ''post-fall syndrome''[11], i.e. excessive fear of falling after a fall. Though fear of falling is indeed related to earlier falls, fear of falling has also been reported in many older people who did not fall at all, suggesting a multi-factorial etiology including other psychological factors such as anxiety and depression [7,12]. Fear of falling has been defined as *a lasting concern about falling that leads to an individual avoiding activities that he/she remains capable of performing* [13]. Fall-related self-efficacy has been used as a proxy for fear of falling and is related to fear of falling but probably with less intensity and emotion [14]. More recently, fear of falling and

Address for correspondence: Jan H. M. Visschedijk, MD, Department of Public Health and Primary Care, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, The Netherlands. Tel: +31-10-2412866. E-mail: j.visschedijk@laurens.nl

fall-related self-efficacy have increasingly been regarded as different concepts [15]. Fall-related self-efficacy focuses particularly on a person's confidence in his or her ability to avoid falling while undertaking activities of daily living (ADL) and is conceptually similar to balance confidence [14]. Though related with fall-related self-efficacy, fear of falling (FoF) can be regarded as a broader concept which includes physiological, behavioral as well as cognitive elements. The distinction between fall related self-efficacy and FoF has also been important when developing and evaluating fall-related psychological measures [16].

The most frequently used scale for fear of falling, which measures different levels of concern about falling, is the Falls Efficacy Scale-International (FES-I), developed and validated by the Prevention of Falls Network Europe (ProFaNE) [17,18]. It is widely used and regarded as a suitable instrument to evaluate for fear of falling among community-dwelling older people [18]. A shorter version has also been developed and tested [19]. Validation studies for the FES-I have been carried out in different patient groups [20–22]. Though there is some minor overlap with patients in these studies, the measurement properties of the FES-I have not been tested in older patients rehabilitating after a hip fracture. Such an evaluation is important since patients with a hip fracture differ from community-dwelling older persons because they have recently experienced a traumatic fall and their general health status is worse, with more disabilities and comorbidity [2]. Therefore, it remains unclear whether the FES-I is a reliable and valid instrument to measure fear of falling after hip fracture. Such an instrument is particularly needed when interventions are designed and implemented to reduce fear of falling in order to improve the outcomes of rehabilitation after hip fracture. A reliable and valid instrument will be useful to select patients for these interventions and to monitor and evaluate outcomes.

Hence, this study aims to evaluate the measurement properties of the FES-I after a hip fracture in patients aged $\geq$ 65 years, based on the guidelines of the COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) group [23]. Therefore, in patients with a hip fracture, we assessed the structural validity, internal consistency, inter-rater reliability, measurement error, construct validity, and floor and ceiling effects of the FES-I.

Methods

Design and study population

For this study we used two study samples of older patients with hip fracture who underwent rehabilitation in a Dutch skilled nursing facility (SNF); in the Netherlands, about 40% of patients with a hip fracture rehabilitate in a SNF [24]. The first study sample consisted of 100 patients who rehabilitated after a hip fracture in 10 different SNFs. Only patients aged $\geq$ 65 years were included. Patients with communication problems and/or who (according to the treating elderly care physician) were unable to respond adequately to questions were excluded. Data collection, which also included information through a questionnaire for the treating elderly care physician and responsible nurse on age, gender, marital status, living situation, comorbidities, complications, short-term and long-term memory, took place between September 2010 and March 2011 [24].

The second study sample consisted of patients who were admitted to the SNF of the PW Janssen Nursing Home in Deventer (the Netherlands). Patients were included between October 2011 and April 2012. These patients were aged $\geq$ 65 years, were admitted because of a hip fracture, and (according to the elderly care physician) were able to answer questions on the FES-I. In the 3rd and 4th week after admission to the SNF all patients were interviewed three times using the FES-I by a psychologist, a physiotherapist and a nurse, after they received a brief collective training on the use of the FES-I. The sequence of the interviewers was randomized and the time between the first and last interview was 10 d or less, with a period of at least 3 d between each interview. Basic information on the participants (age, gender, marital status, living situation and site of the fall) was also collected. In total, 23 patients participated in this part of the study, of whom 21 completed all measurements.

The Medical Ethics Committee of the VU University Medical Center approved the study protocol. All patients gave informed consent for participation.

Measurement instruments

Falls Efficacy Scale-International

Various attempts have been made to assess fear of falling [16]. Single items have been used but generally do not determine the intensity of fear of falling, and do not detect specific changes in fear of falling over time. The Falls Efficacy Scale (FES) was initially developed to solve these problems, focussing on self-confidence not to fall when carrying out certain activities [25]. However, ''self-efficacy'' in performing activities without falling, operationalized in the FES, and actual fear of falling are not the same concepts [26]. Furthermore, the FES suffered from ceiling effects and lacked social activities [18]. Therefore, the Falls Efficacy Scale-International (FES-I) was developed by the ProFaNE group, which also facilitates cross-cultural validation of the instrument for coordinated international studies and comparison. The initial validation was done in English [18], followed by validation in many other languages [27–30].

The FES-I can be completed within 3–4 min. It can be filled in directly by the patient or the information can be collected through an interview, as was done in our study. The FES-I reflects concern about falling when performing 16 ADL. The response to the FES-I consists of four levels ranging from ''not at all concerned'' to ''very concerned'' (score range: 16–64) [26]. The FES-I has shown good measurement properties in community-dwelling older people [26]. In a group of 94 people which were recruited in a postal survey in Germany the Cronbach's alpha was 0.90 and the intra-class correlation was 0.79 [26]. In a sample of 193 participants aged 70 years or more in the Netherlands these figures were, respectively, 0.96 and 0.82 [26].

One-item fear of falling instrument

The one-item fear of falling instrument poses one question: *Are you afraid of falling?* It has four answer options ''not at all'', ''a little'', ''quite a bit'', and ''very much'' [16]. The test–retest coefficient kappa was 0.66 with a retest after 4–7 d [16]. Although it is often used, the evidence for adequate validity of one-item instruments is weak [16]. However, when considering that the FES-I measures fall-related self-efficacy, researchers have been advised to add a single-item measure specific to fear of falling to ensure measurement of both concepts [31]. Information for the one-item fear of falling instrument was also collected by interview.

Instruments for psychological and cognitive factors

Data related to psychological constructs were collected through interviews with the participants by an elderly care physician or psychologist. No data on the measurement properties specific for patients with hip fractures of these patient-reported outcomes were found in the literature.

Depressive symptoms were measured using the Geriatric Depression Scale 8-item version (GDS8); this is an adaptation of the GDS30 that better fits institutionalized older people [32].

The GDS8 has eight items (score range: 0–8) with higher scores indicating more depression. The GDS has good measurement properties; it was validated using the DSM-IV diagnosis for depression, is internally consistent (Cronbach's alpha = 0.80) and has high sensitivity rates for major (96.3%) and minor (83.0%) depression [32].

Anxiety was assessed using the anxiety component of the Hospital Anxiety and Depression Scale (HADS-A) [33]. The HADS-A has seven items (score range: 0–21) with higher scores indicating more anxiety. The measurement properties of the HADS are good. Cronbach's alpha for the HADS-A ranges from 0.68 to 0.93 and the validity is good when compared with other commonly used questionnaires [34].

Self-efficacy was measured using the Self-Efficacy Scale (SES) [35]. This scale has 10 items (score range: 0–30) with a higher score indicating a higher competence to cope with different challenges. The scale has been used in numerous studies and generally yielded internal consistency (alpha: 0.75–0.91). For a total sample of 19 120 respondents from 25 countries Cronbach's alpha was 0.86. The scale can be regarded as a uni-dimensional instrument [36]. The test–retest score is fair and, for example, was reported to be $r = 0.67$ in German cardiac surgery patients [37]. Evidence for the validity of the SES has also been published [36].

Impairment in short- and long-term memory was rated based on an assessment by the responsible nurse using the Cognitive Performance Scale from the Minimum Data Set of the nursing home resident assessment instrument [38].

Functional outcomes

Three functional outcome measurements were used to measure balance and walking ability. Both the Performance-Oriented Mobility Assessment (POMA) and the Timed-Up-and-Go (TUG) test measure balance and walking ability, while functional ambulation categories (FAC) only give an indication of a patient's walking ability. With the POMA, the participant follows the instructions of the physiotherapist, who scores the different components of the test. The score of the POMA ranges from 0 to 28, with a higher score indicating better balance and walking ability [39]. The inter-rater and test–retest reliability for the POMA is excellent ($r = 0.82$–0.93) [39]. The correlation with reference performance tests ($r = 0.65$–0.70) indicates satisfactory construct validity for the POMA [39].

With the TUG the physiotherapist measures the time it takes to stand up from a chair, walk 3 m, turn around, and walk back to the chair and sit down, all at a comfortable speed [40]. The inter-rater and intra-rater reliability is high and the construct validity is reported to be fair when compared with other measures that assess walking ability and balance in community-dwelling older people [41].

The FAC was scored by the physiotherapist. The score of the FAC ranges from 0 to 5, with higher scores indicating a person's ability to walk more independently [42]. The inter-rater reliability of the FAC is high ($r = 0.91$) and the FAC has a good construct validity in relation to other tests such as the 6-minute walking test and walking velocity [42].

ADL after hip fracture were measured using the Barthel Index (BI) [43]. The BI was scored by the responsible nurse. It has 10 items and assesses the degree of support a person needs in performing ADL, such as eating, getting dressed and going to the toilet. Although the index initially focused on stroke patients, it is used for a wide variety of patients. The score of the BI ranges from 0 to 20, with a higher score indicating more independence in ADL activities. The internal consistency of the score is high; for example, it is 0.84 in patients with a stroke [44]. The inter-rater reliability is also high ($r = 0.88$–0.99) [45] and the BI has proven to be a valid measure for ADL [46].

A fall was defined as an event that results in a person coming to rest inadvertently on the ground or lower level [47]. It includes also falls from internal causes such as fainting or collapse. Besides the site of the fall, indoor versus outdoors, the fall history of the participants was assessed. Fall history was measured on a 3-point scale by asking the participants how often they had fallen during the last 6 months before hip fracture. The answer categories were: not at all, one time, or more than one time.

Assessment of measurement properties

Structural validity

Structural validity is defined as ''the degree to which the scores of a measurement instrument are an adequate reflection of the dimensionality of the construct to be measured'' [15] and can be assessed by factor analysis.

Internal consistency

Internal consistency is the interrelatedness among the items in a scale [23]. Different items in an instrument may ask the same questions in a slightly different manner to reliably capture the respondent's opinion or level of function. The Cronbach's alpha is considered to be an adequate measure of internal consistency when it is shown that the scale is uni-dimensional (e.g. by factor analysis). A low Cronbach's alpha indicates a lack of correlation between the items in a scale, which implies that summarizing the items is unjustified. A very high Cronbach's alpha (> 0.95) reflects high correlations among the items in the scale, which may indicate redundancy of one or more items [48].

Reliability

Reliability is the proportion of the total variance in the measurement that is due to true differences between patients. This refers to the degree to which the measurement instrument is free from measurement error, and estimates the extent to which scores for patients who have not changed are the same for repeated measurements, e.g. by different raters (inter-rater reliability) [23].

Measurement error

Measurement error is the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured [23]. Measurement error can be expressed as the standard error of measurement (SEM) or the smallest detectable change (SDC). These calculations are expressed in the unit of measurement of the scale of the instrument. The SEM represents the standard deviation (SD) of repeated measures of one patient. The SDC represents the minimal change that a patient has to show on the scale to ensure that the observed change is real and not just an inter-rater measurement error.

Construct validity

Validity is the degree to which an instrument measures the construct it is supposed to measure. In the absence of a gold standard, as is the case for the FES-I, construct validity refers to the extent to which a particular measure relates to other measures based on theoretically derived hypotheses for the constructs that are being measured. We used the one-item fear of falling instrument, the HADS-A, the GDS8, the SES, the POMA, the TUG, the FAC score, the BI, and the fall history, including both falls indoors and outdoors, to assess the construct validity of the FES-I for patients with a hip fracture. Based on our knowledge at the time of design of the study, we formulated 11 ''*a priori*'' hypotheses for the minimal level of validity. We expected the FES-I to have the highest correlation with the one-item fear of falling instrument, because both measure a similar construct (correlation of > 0.50). Also, the HADS-A was expected to be highly associated with FES-I because of the similarities of both constructs; we expected a correlation of 0.30–0.50 between these constructs. The FES-I was expected to have a higher correlation with the HADS-A than with the GDS and SES, since these constructs are substantially different; we expected a correlation of $\leq$0.30 between the FES-I and the GDS8, and between the FES-I and the SES. Furthermore, we expected a smaller correlation with functional outcomes such as the POMA, the TUG, the FAC score, BI and fall history (correlation of ≤ 0.30). In the case that $\geq 75\%$ of the hypotheses can be confirmed, the construct validity is considered to be adequate [49].

Floor and ceiling effects

The presence of floor or ceiling effects may have a negative effect on the quality of the instrument. If a group of patients scores mainly in the extremes or within the SDC of the extremes, the responsiveness may be limited.

Statistical analyses

We first assessed structural validity to evaluate whether the scale is uni-dimensional. Confirmatory factor analysis for categorical items was performed in Mplus (Meuthen and Meuthen, Los Angeles, CA) using weighted least squares, with means and variance adjustment. We examined factor loadings and model fit. Factor loadings represent the correlation between the items of the FES-I and the factor (the underlying dimensions). Analogous to Spearson *r*, the squared factor loading is the percentage of variance in the indicator variable explained by the factor. Factor loadings are generally considered to be meaningful when they are $\geq$0.30 or 0.40 [50]. We considered factor loadings of $\geq$0.50 to be appropriate. The Comparative Fit Index (CFI), the Tucker–Lewis Index (TLI), and the root mean square error of approximation (RMSEA) were used as measures for model fit. A CFI and TLI $\geq$0.95 and a RMSEA of $\leq$0.05 were considered to be an adequate fit. For a moderate fit, values of $\geq$0.90 and $\leq$0.08 were used. The internal consistency was assessed by calculating Cronbach's alpha, using the widely accepted cut-off of $\geq$0.7 [49].

Reliability was assessed by calculating the intra-class correlation coefficient (ICC) with a 95% confidence interval (95% CI). A two-way mixed-effects model for absolute agreement was used. An ICC $\geq$0.7 was considered to be good [51]. The SEM was calculated from the square root of the variance between the raters and the error variance of the ICC. The SDC was calculated as $1.96 \times \sqrt{2} \times \text{SEM}$. Because most variables were not evenly distributed validity was tested by calculating Spearman correlation coefficients.

We calculated the floor and ceiling effects as the percentage of the participants who had the minimum and maximum score, respectively (i.e. 16 or 64, respectively). Floor or ceiling effects were considered to be present when $\geq$15% of the respondents achieved the minimum or maximum possible score [52]. Analyses were performed with SPSS for Windows (Version 19, SPSS, Inc., Chicago, IL).

Results

Table 1 presents the characteristics of the two study groups. In Group 1 and Group 2 the mean FES-I was 32.2 and 36.0, and the mean age was 83.1 and 83.2 years, respectively. In both groups the majority of the participants were widows and lived alone. Most falls, resulting in a hip fracture occurred indoors. Only 30% of the

Table 1. Characteristics of the two study groups.

	Group 1 ($n = 100$)	Group 2 ($n = 21$)
FES-I, mean (SD)	32.2 (9.6)	36.0 (10.9)
Age in years, mean (SD)	83.1 (8.3)	83.2 (7.2)
Female, *n* (%)	75 (75%)	19 (90%)
Marital status, *n* (%)		
Married	18 (18%)	3 (14%)
Widow/widower	68 (68%)	16 (76%)
Divorced	4 (4%)	1 (5%)
Single	10 (10%)	1 (5%)
Living alone, *n* (%)	78 (78%)	17 (81%)
Site of fall[a], *n* (%)		
Indoors	70 (70%)	17 (81%)
Outdoors	30 (30%)	4 (19%)
Fall history (No of falls in half year before hip fracture)		
Nil	77 (77%)	
Once	11 (11%)	
Twice or more	12 (12%)	
ADL (BI), mean (SD)	12,7 (4,6)	
Ability to walk independent[b]	44 (45%)	
TUG, mean (SD)	38.7 (31.7)	
POMA, mean (SD)	17.0 (6.3)	
Number of comorbidities, mean (SD)	3.5 (1.5)	
Number of complications, mean (SD)	1.6 (1.4)	
Impairment of short-term memory, number (%)	19 (19%)	
Impairment of long-term memory, number (%)	6 (6%)	

FES-I, Falls Efficacy Scale-International (range 16–64); SD, standard deviation. Nr, number; ADL, activities of daily living; BI, Barthel Index (range 0–20); TUG, timed up and go test; POMA, performance oriented mobility assessment (range 0–28).

[a]This refers to the place where the participant fell when fracturing the hip.

[b]This refers to a FAC score of 4 of 5.

participants in group 1 and 19% of the participants in group 2 fell outdoors.

Structural validity

Table 2 presents the results of the confirmatory factor analysis on the baseline data. A 1-factor model fitted the data adequately. The CFI was 0.994, the TLI was 0.993, and the RMSEA was 0.047. No items had a factor loading $\leq$0.50 and only two items had a factor loading $\leq$0.70, i.e. item 2 (loading 0.695) and item 4 (loading 0.669). Thus, there is strong evidence for the uni-dimensionality of the FES-I.

Internal consistency

The Cronbach's alpha was 0.94, which implies good internal consistency.

Reliability

The ICC for all raters was 0.72 (95% CI: 0.52–0.87). The ICCs for the physiotherapist versus the nurse, the physiotherapist versus the psychologist, and the nurse versus the psychologist were 0.70 (95% CI: 0.41–0.87), 0.78 (95% CI: 0.53–0.90) and 0.69 (95% CI: 0.34–0.87), respectively. The SEM for all raters was 6.4 and the SDC was 17.7. Table 3 presents the mean scores of the physiotherapist, nurse and psychologist.

Construct validity

Construct validity was assessed by testing the ''*a priori*''-defined hypotheses. Correlations between the FES-I and the other constructs are presented in Table 4. The table shows that hypothesis numbers 1, 3, 5 and 11 could be confirmed; this is 36% of all the hypotheses.

Floor and ceiling effects

There were no floor or ceiling effects: 0% of all patients had the maximum score (64) and 1% had the minimum score (16). When assessing how many participants had a score within the SDC (17.7) of the maximum (i.e. 47 or higher), i.e. indicating a high level of fear of falling, the percentage was 8%. For the minimum score (i.e. 33 or lower), i.e. indicating a low level of fear of falling, the percentage was 54%.

Table 2. Factor loadings of the Falls-Efficacy Scale-International.

Item/Factor		Estimate	Standard error
F1	Cleaning the house	0.826	0.035
F2	Getting dressed /undressed	0.695	0.073
F3	Preparing simple meals	0.796	0.045
F4	Taking a bath or shower	0.669	0.055
F5	Going to the shop	0.910	0.025
F6	Getting in or out of a chair	0.842	0.035
F7	Going up or down stairs	0.744	0.049
F8	Walking around outside	0.831	0.036
F9	Reaching up or bending down	0.782	0.042
F10	Answering the telephone	0.729	0.051
F11	Walking on a slippery surface	0.765	0.047
F12	Visiting a friend/relative	0.876	0.032
F13	Going to a place with crowds	0.807	0.040
F14	Walking on an uneven surface	0.835	0.033
F15	Walking up or down a slope	0.834	0.037
F16	Going out to a social event	0.955	0.018

Discussion

This study shows that the FES-I is an internally consistent and reliable instrument to measure fear of falling in patients after a hip fracture. For this population the instrument is uni-dimensional; it has no floor and ceiling effects. Based on our ''*a priori*'' hypotheses the validity is fair but not excellent, since we could confirm only 4 of the 11 predetermined hypotheses. When testing the construct validity, the correlation with the one-item instrument for fear of falling was strong ($r = 0.68$) and higher than that in a recent study performed in China ($r = 0.42$) [29]. Also, the FES-I was found to have a stronger relation with physical performance constructs (such as mobility, balance and ADL) than with psychological constructs (such as anxiety and self-efficacy).

Others also found a strong correlation with physical performance constructs, such as the TUG. For example, a study in Greece reported the Pearson correlation to be 0.638 [28]. In a recent validation of the Chinese version of the FES-I among 399 community-dwelling Chinese older people, the FES-I score was significantly higher in participants with poor physical performance [29]. In this Chinese study, Pearson correlations between the FES-I and the TUG, the IADL and depressive symptoms were 0.22, 0.21 and 0.13, respectively; this also indicates a better relation with physical performance than with psychological factors. Similarly, in a validation study in Turkey among 70 older people, the Pearson correlation coefficient between the FES-I and the Modified Barthel Index and TUG was 0.622 and 0.743, respectively [30]. In fact, our ''*a priori*''-defined hypotheses related to physical performance and psychological concepts were not in line with more recent studies. We also found that the FES-I in older patients with a hip fracture is much stronger correlated with physical performance than with psychological factors such as anxiety. It also emphasizes that fall-related self-efficacy and fear of falling are related but different concepts.

In our study the Spearmen correlation coefficient between the FES-I and fall history was only 0.17; this is much lower than in a study among persons with multiple sclerosis ($r = 0.46$) [53]. Fall history reflects in our study the number of falls over the last 6 months in addition to the fall in which the participant fractured his or her hip. As a result, a group of non-fallers did not exist in our study and all participants experienced at least one traumatic fall with tremendous consequences, such as long-lasting pain, admission to a hospital, surgical repair and inability to walk. This may have weakened the relation between fall history and fear of falling.

Our factor analysis suggested uni-dimensionality of the FES-I. In other studies among community-dwelling older persons, the factor analysis was suggestive for two underlying factors, i.e. concern about falling during ADL, and concern about falling during social activities [27,29]. It is possible that, after a hip fracture, rehabilitating older patients are mainly concerned with basic ADL and hardly discriminate between these activities and social activities (which may seem less relevant to them during rehabilitation).

Our Cronbach's alpha of 0.94 indicates a good internal consistency and is similar to studies among community-dwelling elderly in Brazil [27] and China [29], as well as among other patient groups such as cognitively impaired geriatric patients [21] and patients with multiple sclerosis [53]; in these latter studies the Cronbach's alpha were 0.93, 0.94, 0.93–0.95 and 0.94, respectively. In a study by Kempen et al. among community-dwelling older persons in Germany, the Netherlands and the UK, the Cronbach's alpha was 0.90, 0.96 and 0.97, respectively [26]. Our inter-rater reliability was ICC = 0.72 which is good. A higher reliability, e.g. in the studies of Camargos et al. (ICC = 0.91) [27] and Yardley et al. (test–retest reliability ICC = 0.96) [17], has

Table 3. Inter-raters reliability of the Falls Efficacy Scale-International.

Mean (SD)			SEM	SDC	ICC (95% CI)
Observer 1 (physiotherapist)	Observer 2(nurse)	Observer 3 (psychologist)			
36.3 (11.3)	33.5 (11.9)	38.3 (12.5)	6.4	17.7	0.72 (0.52–0.87)

SD, standard deviation; SEM, standard error of measurement; SDC, smallest detectable change; ICC, intra-class correlation coefficient.

Table 4. Validity of the Falls Efficacy Scale-International (FES-I).

No.	Hypothesis	Comparison measurement instrument	Observed correlation with FES-I (Spearman correlation)	Hypothesis confirmed
1	A correlation of > 0.5 was expected between the FES-I and the one-item fear of falling instrument	One-item fear of falling instrument	0.68	Yes
2	A correlation of > 0.3 but ≤ 0.5 was expected between the FES-I and the HADS/A	HADS/A	0.30	No
3	The correlation between the FES-I and the HADS/A is stronger than that between the FES-I and the GDS8	HADS/A	0.30	Yes
		GDS8	0.03	
4	The correlation between the FES-I and the HADS/A is stronger than that between the FES-I and the SES	HADS/A	0.30	No
		SES	−0.32	
5	A correlation of ≤ 0.3 was expected between the FES-I and the GDS8	GDS8	0.03	Yes
6	A correlation of ≤ 0.3 was expected between the FES-I and the SES	SES	−0.32	No
7	A correlation of ≤ 0.3 was expected between the FES-I and the POMA	POMA	−0.43	No
8	A correlation of ≤ 0.3 was expected between the FES-I and the TUG	TUG	0.31	No
9	A correlation of ≤ 0.3 was expected between the FES-I and FAC score after fracture	FAC score after fracture	0.31	No
10	A correlation of ≤ 0.3 was expected between the FES-I and ADL after fracture	ADL after fracture	−0.34	No
11	A correlation of ≤ 0.3 was expected between the FES-I and the fall history	Fall history	0.17	Yes

HADS/A, Hospital Anxiety and Depression Scale/Anxiety component; GDS8, Geriatric Depression Scale-8 item version; SES, Self-Efficacy Scale; POMA, performance oriented mobility assessment; TUG, timed up and go test; FAC, functional ambulation categories; ADL, activities of daily living.

been reported. In a study in China among community-dwelling older people the inter-rater reliability was very high (ICC = 0.95) [29]. In the study of Kempen et al., which included community-dwelling older persons in the Netherlands, the test–retest reliability was also higher (ICC = 0.82) [26]. Reasons for the lower correlation coefficient in our study might be because: (i) the relatively older and vulnerable patients (some with a cognitive disorder, most with rather high number of comorbidities) may have been less consistent in answering the FES-1 questions, and (ii) different types of professionals rated the FES-I.

The absence of floor and ceiling effects is common in most studies on the FES-I [28]. In a study among cognitively impaired patients, the floor effect (minimum score) was 3.2% and the ceiling effect (maximum score) was 0% [54]. In our study the SDC was substantial (i.e. 17.7 compared to a range of 16–64). Though this may make it more difficult to measure changes in fear of falling in patients with a low level of fear of falling, since 54% of the participants had a score of ≤ 33, for interventions which are targeted towards patients with higher levels of FoF, improvements can be correctly measured.

Since the FES-I particularly focuses on fall-related self-efficacy and does not cover all elements of fear of falling it has been advised to use simultaneously a one 1-item instrument in research [31]. This will ensure that besides the concept of fall-related self-efficacy also the concept of fear of falling is measured including more emotional and physiological dimensions of fear of falling. Recently also a modification of the FES was made for nursing homes, i.e. the Nursing Home Falls Self-Efficacy Scale [20]. This instrument has items on both self-efficacy expectations and outcome expectancy, focussing on the consequences of falling (embarrassment, pain, risk of fracture, etc.). More research is required to assess whether this instrument can also be relevant for older patients rehabilitating in a SNF of a nursing home.

In some studies the FES-I was administered through self-reporting [26]. In a study by Hauer et al. to validate the FES-I in geriatric patients, the FES-I was administered by both self-report and interview-based questionnaires [54] The ICCs were respectively for the interview and self-reported method 0.744 and 0.584. The authors concluded that in vulnerable older persons, especially with cognitive impairment, an interview-based method is recommended. We also used the interview-based method, which may have had a positive influence on the outcomes of the measurement properties. However, since 19% and 6% of the participants had, respectively, short-term and long-term cognitive impairments, the answers to the FES-I may have been less consistent, hampering the reliability of the FES-I, even when using an interview-based approach. In addition, since the FES-I particularly measures concerns about falling, participants with impaired cognition may evaluate their risk to fall different from those who have no cognitive impairment. More research is needed to assess how strong the impact of such conditions is on fall-related self-efficacy.

Strength of our study is that we assessed the measurement properties of the FES-I in a population of vulnerable people in which fear of falling may have substantial consequences for daily activities and quality of life [6]. To our knowledge, this is the first study to assess the measurement properties of the FES-I in people aged $\geq$ 65 years who had a traumatic fall resulting in a hip fracture. In addition, to assess the validity of the FES-I, we included a wide variety of tests used in daily practice. Although the concept of fear of falling needs further research, the FES-I seems to be a suitable instrument to assess FoF among patients after a hip fracture. Nevertheless, future studies need to further explore this concept, particularly with regard to how it interacts with other concepts of psychological and physical performance.

Conclusion

The results of the present analysis indicate that the reliability and structural validity of the FES-I in patients aged $\geq$ 65 years after a hip fracture is good. When assessing the construct validity of the FES-I, the construct seems to be more closely related to functional constructs than to psychological constructs. This may indicate that the concept measured by the FES-I does not capture all aspects of fear of falling.

Acknowledgements

The authors thank all the participants in the study, the physiotherapist, nurse and psychologist who interviewed the participants using the FES-I, the participating nursing homes, and the university nursing home networks of the Free University of Amsterdam (UNO-VUmc) and University Medical Center Leiden (UNC-ZH).

Declaration of interest

The authors report no conflict of interest.

References

1. Gullberg B, Johnell O, Kanis JA. World-wide projections for hip fracture. Osteoporos Int 1997;7:407–13.
2. Leibson CL, Tosteson AN, Gabriel SE, et al. Mortality, disability, and nursing home use for persons with and without hip fracture: a population-based study. J Am Geriatr Soc 2002;50:1644–50.
3. Balen van R, Steyerberg EW, Polder JJ, et al. Hip fracture in elderly patients: outcomes for function, quality of life, and type of residence. Clin Orthop Relat Res 2001;390:232–43.
4. Oude Voshaar RC, Banerjee S, Horan M, et al. Fear of falling more important than pain and depression for functional recovery after surgery for hip fracture in older people. Psychol Med 2006;36: 1635–45.
5. Mossey JM, Mutran E, Knott K, Craik R. Determinants of recovery 12 months after hip fracture: the importance of psychosocial factors. Am J Public Health 1989;79:279–86.
6. Jellesmark A, Herling SF, Egerod I, Beyer N. Fear of falling and changed functional ability following hip fracture among community-dwelling elderly people: an explanatory sequential mixed study. Disabil Rehabil 2012;34:2124–31.
7. Scheffer AC, Schuurmans MJ, Dijk van N, et al. Fear of falling: measurement strategy, prevalence, risk factors and consequences among older people. Age Ageing 2008;37:19–24.
8. Muche R, Eichner B, Gebhard F, et al. Risikofaktoren und prognosemöglichkeiten für mortalität und soziofunktionelle Einschränkungen bei Älteren nach proximalen Femurfrakturen. Euro J Ger 2003:5:187–94.
9. Ingemarsson AH, Frändin K, Hellström K, Rundgren A. Balance function and fall-related efficacy in patients with newly operated hip fracture. Clin Rehabil 2000;14:497–505.
10. Visschedijk J, van Balen R, Hertogh C, Achterberg W. Fear of falling in patients with hip fractures: prevalence and related psychological factors. J Am Med Dir Assoc 2013:14;218–20.
11. Murphy J, Isaacs B. The post-fall syndrome. A study of 36 elderly patients. Gerontology 1982;28:265–70.
12. Legters K. Fear of falling. Phys Ther 2002;82:264–72.
13. Tinetti ME, Powell L. Fear of falling and low self-efficacy: a case of dependence in elderly persons. J Gerontology 1993;48:35–8.
14. Hadjistavropoulos T, Delbaere K, Fitzgerald TD. Reconceptualizing the role of fear of falling and balance confidence in fall risk. J Aging Health 2011;23:3–23.
15. Moore DS, Ellis R. Measurement of fall-related psychological constructs among independent-living older adults: a review of the research literature. Aging Ment Health 2008;12:684–99.
16. Jørstad EC, Hauer K, Becker C, Lamb SE. Measuring the psychological outcomes of falling: a systematic review. J Am Geriatr Soc 2005;53:501–10.
17. Yardley L, Beyer N, Hauer K, et al. Development and initial validation of the Falls Efficacy Scale-International (FES-I). Age Ageing 2005;34:614–19.
18. Delbaere K, Close JCT, Mikolaizak AS, et al. The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. Age Ageing 2010;39:210–16.
19. Kempen GI, Yardley L, van Haastregt JC, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 2008;37:45–50.
20. Lach HW, Ball LJ, Birge SJ. The Nursing Home Falls Self-Efficacy Scale: development and testing. Clin Nurs Res 2012;21:79–91.
21. Hauer K, Yardley L, Beyer N, et al. Validation of the Falls Efficacy Scale and Falls Efficacy Scale International in geriatric patients with and without cognitive impairment: results of self-report and interview-based questionnaires. Gerontology 2010;56:190–9.
22. Denkinger MD, Igl W, Coll-Planas L, et al. Practicality, validity and sensitivity to change of fear of falling self-report in hospitalised elderly – a comparison of four instruments. Age Ageing 2009;38: 108–12.
23. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63;737–45.
24. Visschedijk JHM, Caljouw MAA, van Balen R, et al. Fear of falling after hip fracture in vulnerable older persons rehabilitating in a skilled nursing facility. J Rehabil Med 2014;46:258–63.
25. Tinetti ME, Richman D, Powel L. Falls efficacy as a measure of fear of falling. J Gerontol 1990;45:239–43.
26. Kempen GI, Todd CJ, Van Haastregt JC, et al. Cross-cultural validation of the Falls Efficacy Scale International (FES-I) in older people: results from Germany, the Netherlands and the UK were satisfactory. Disabil Rehabil 2007;29:155–62.
27. Camargos FFO, Dias RC, Dias JMD, Freire MTF. Cross-cultural adaptation and evaluation of the psychometric properties of the Falls Efficacy Scale – International among elderly Brazilians. Rev Bras Fisioter 2010;14:237–43.
28. Billis E, Strimpakos N, Kapreli E, et al. Cross-cultural validation of the Falls Efficacy Scale International (FES-I) in Greek community-dwelling older adults. Disabil Rehabil 2011;33:1776–84.
29. Kwan MMS, Tsang WWN, Close JCT, Lord SR. Development and validation of a Chinese version of the Falls Efficacy Scale International. Arch Gerontol Geriatr 2013; 56:169–74.
30. Ulus Y, Durmus D, Akyol Y, et al. Reliability and validity of the Turkish version of the Falls Efficacy Scale International (FES-I) in community-dwelling older persons. Arch Gerontol Geriatr 2012;54: 429–33.
31. Greenberg S. Analysis of measurement tools for fear of falling for high-risk, community dwelling older adults. Clin Nurs Res 2012;21: 113–30.
32. Jongenelis K, Gerritsen DL, Pot AM, et al. Construction and validation of a patient- and user-friendly nursing home version of the Geriatric Depression Scale. Int J Geriatr Psychiatr 2007;22:837–42.
33. Spinhoven P, Ormel J, Sloekers PP, et al. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med 1997;27:363–70.

34. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002;52:69–77.
35. Teeuw B, Schwarzer R, Jerusalem M. Dutch adaptation of the general self-efficacy scale; 1994. Available from: http://userpage.fu-berlin.de/~health/dutch.htm [last accessed 2 Oct 2012].
36. Scholz U, Gutiérrez Dona B, Sud S, Schwarzer R. Is general self-efficacy a universal construct. psychometric findings from 25 countries. Eur J Psychol Assess 2002;18:242–51.
37. Schröder KEE, Schwarzer R, Konertz W. Coping as a mediator in recovery from cardiac surgery. Psychol Health 1998;13:83–97.
38. Gruber-Baldini AL, Zimmerman SI, Mortimore E, Maganizer J. The validity of the minimum data set in measuring the cognitive impairment of persons admitted to nursing homes. J Am Geriatric Soc 2000;48:1601–6.
39. Faber MJ, Boscher RJ, van Wieringen PC. Clinimetric properties of the performance oriented mobility assessment. Phys Ther 2006;86: 944–54.
40. Podsiadlo D, Richardson S. The timed 'Up and Go': a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39:142–8.
41. Lin MR, Hwang HF, Hu MH, et al. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. J Am Geriatr Soc 2004;52:1343–8.
42. Mehrholz J, Wagner K, Rutte K, et al. Predictive validity and responsiveness of the functional ambulation category in hemiparetic patients after stroke. Arch Phys Med Rehabil 2007;88:1314–19.
43. Mahony FI, Barthel DW. Functional evaluation: the Barthel Index. Md Stat Med J 1965;14:61–5.
44. Hsueh IP, Lin JP, Jeng JS, Hseih CL. Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke. J Neurol Neurosurg Psychiatr 2002; 73:188–90.
45. Roy CW, Togneri J, Hay E, Pentland B. An interrater reliability study of the Barthel Index. Int J Rehabil Res 1988;11: 67–70.
46. Quinn TJ, Langhorne P, Stott DJ. Barthel Index for stroke trials: development, properties, and application. Stroke 2011;42:1146–51.
47. WHO global report on falls prevention in older age. Geneva: World Health Organization; 2007.
48. Pynsent PB. Choosing an outcome measure. J Bone Joint Surg Br 2001;83:792–4.
49. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
50. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. Psychol Assess 1995; 7:286–99.
51. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013;22:1889–905.
52. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate?. Qual Life Res 1995;4:293–307.
53. Vliet van R, Hoang P, Lord S, et al. Falls efficacy scale-international: a cross-sectional validation in people with multiple sclerosis. Arch Phys Med Rehabil 2013;94:883–9.
54. Hauer KA, Kempen GIJM, Schwenk M, et al. Validity and sensitivity to change of the Falls Efficacy Scales International to assess fear of falling in older adults with and without cognitive impairment. Gerontology 2011;57:462–72.